AIRS Product Development

The AIRS Software Development System

Steven Friedman
AIRS Science Processing

April 26, 2011

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration. © 2011 California Institute of Technology. Government sponsorship acknowledged.
Topics

- AIRS Software Development Activities: V6 Specific
  - CM Processes
  - Development and Testing

- AIRS Computing Facility Status

- Impact of System Intrusion on AIRS Activities

- V6 Completion Phasing
AIRS Software Development Activities
- CM Processes

• AIRS uses the same successful software development process it started with:
  • Science Team develops code
  • JPL integrates, builds and tests code
  • JPL delivers “production code” to GES DISC

• Code packages are configuration managed
  - Harvest CM System
  • One change per package
  • Packages can be backed-out from the configuration
  • Several packages typically collected into a build
  • Each build is tested/evaluated against standard test dataset
  • Significant builds are “verified” or “validated” with extended test sets and correlative data
• Version 6 development began on 2007-07-11 w/ V5.1.0.0
  • Since then more than 45 builds
  • Several significant tests conducted to assess our progress
  • Current V6 development version is V5.7.5.0

- V5.1.7.0 - 2007-10-02 - MW RTA an tuning
- V5.3.0.0 - 2008-02-27 - AMSU-A Ch4 quick-fix (still using)
- V5.3.2.0 - 2008-09-20 - Surface emissivity hinge-points
- V5.4.0.0 - 2009-01-26 - Improved CO2, surface retrieval
- V5.4.5.0 - 2009-05-20 - ECMWF as climatology startup opt.
- V5.4.14.0 - 2009-12-14 - Remove AMSU-A Ch 4,5 from MW and final retrievals
- V5.5.1.0 - 2010-02-25 - Incorporate SCCN as startup option
- V5.6.2.0 - 2010-07-19 - Stability parameters
- V5.6.3.0 - 2010-07-17 - New Regression
- V5.7.0.0 - 2010-10-07 - MODIS Emissivity, New var. freq. RTA, cloud phase

ALL MAJOR V6 COMPONENTS!
AIRS Software Development Activities  
- Testing Status

- A few more builds, getting us to another benchmark test:  
  - Testing Start-up options...
    - Cloudy-Regression (tried and true)
    - Climatology (AIRS Climatology)
    - SCCN
      - ECMWF as startup previously dismissed

- Large test dataset produced
  - Focus Days (24-day cycle) for full mission
  - RaObs Match-ups every 5-days for most of mission
  - Special topical Match-ups: Desert and GPS

- Start-up of these efforts affected by system intrusion  
  - Most products produced by April 1
The AIRS Team Leader Science Computing Facility (TLSCF) is aging and reaching “End of Service Life”

- Maintaining the TLSCF beyond this point is not viable without an upgrade... too expensive!
- Previous system replacement took place more than 5 years ago
- It is time to rebuild the TLSCF for the next 5-years.

TLSCF System Replacement Plan formulated
- Leverages technology improvements since last upgrade
- More compute power and storage
- No additional energy footprint (electricity and cooling)

For hardware, this is a natural evolutionary process!
AIRS Computing Facility Status -
Historical Perspective

• Pre-launch configuration:
  • Monolithic SGI processors (2) - approx. 2.2x capability
  • Data archive 500 GB disk archive upgraded to 1.5 TB
    all with tape backup for data not on disk

• Current configuration:
  • TLSCF “cluster” 20 dual processors running Linux (5-10x)
  • Supported and appended to by many quad and eight-CPU
    systems, enhancement to “science computers”
  • Data archive is now 120 TB and soon to be 200+ TB
    still with tape backup for data not on disk
    but... not all data can be housed in the tape enclosure
• **Future Configuration**
  - Multiple 16+ processors per server unit
  - Data archive consisting of 200 - 500 TB disk plus next-generation tape archive system inevitably, not all tapes will fit in the tape enclosure

• **First system components have been replaced this year**
  - Other hardware is in-house but not installed due to V6 activity
  - Additional replacements in key systems planned soon

• **Goal is to replace necessary systems without:**
  - Service interruptions
  - Extending the AIRS version improvement schedule
  - Incurring excessive costs overall and in any given year
AIRS/PEATE System Intrusion

- Hackers broke into AIRS email server on December 12, 2010 - gaining access via recently identified vulnerability in server

- 2010-12-13 - Intrusion detected by AIRS System Administrator
  - Affected systems blocked from internet - **good and bad aspects**
    - blocks hacker access
    - also cripples desired external network traffic
    - allows local processing to continue
  - 25 servers/workstations affected

- JPL Computer Security and NASA Inspector General’s Office begin investigation

- 2010-12-14 to 2011-01-03 - forensic analysis
  - Systems pulled 1 by 1 by JPL Computer Security and NASA
AIRS/PEATE System Intrusion (cont’d.)

2010-12-27 to 2011-01-04 Systems are released after forensics (one at a time)

2010-01-07 All affected systems taken offline

2011-01-07 to 2011-01-12 Restoration of critical systems

2011-02-16 Last affected system restored to operation

2011-02-22 to 2011-03-16 Restoration of data archive

2011-02-22 Recommenced data processing operations
AIRS/PEATE System Intrusion

• Impacted both AIRS and NPP PEATE Projects
  • 6-week schedule slip for PEATE
    - AIRS schedule impact TBD, but at least 6 weeks
  • Individual work efforts hampered
  • More than 900 work hours expended to rebuild the system!
  • Impact assessment for NASA IG under way
  • Net impact of intrusion estimated to be more than $585,000

• In the end, since intrusion was almost immediately detected
  • No loss/corruption of data or source code!
  • Hackers left back-door entry points, SPAMBots, loggers
    - all removed during restoration
  • The original point of access was removed with upgrade patch
  • Fast action saves the day!
We’ve been working on V6 for quite some time...

We are close to making some major decisions

and

We are close to moving beyond V6

There is still much to be accomplished!
Remaining V6 Goals

- We may have decided which start-up state to use...
  
  *but other V6 decisions still needed to be made.*

  - Further improvement of boundary layer retrievals (?)
  - Refinement of QC and other status information
  - Incorporation of back-end features (cloud properties, trace gases, ...)
  - Level 3, possibly new Level 3 climate products
V6 Milestones

• **AIRS Project Goals**
  - V6 must meet AIRS Project’s and Science Team’s goals
  - Deliver V6 to GES DISC before end of FY11

• **To accomplish this task...**
  - Concur on which start-up state to use
  - Determine whether Key V6 goals have been attained
    - agree to fix what we can
    - possibly defer some goals to V7

• **Milestones...**
  - June 30 - L2 must be completed
  - August 1 - Testing ends
  - August 23 - Documentation and code to GES DISC
Backup Materials
<table>
<thead>
<tr>
<th>Task Name</th>
<th>Start</th>
<th>Finish</th>
<th>Duration</th>
<th>% Complete</th>
<th>Predecessor</th>
</tr>
</thead>
<tbody>
<tr>
<td>14 - System Intrusion (Unplanned)</td>
<td>Sun 12/12/10</td>
<td>Wed 3/16/11</td>
<td>64 days</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>15 - Intrusion</td>
<td>Sun 12/12/10</td>
<td>Sun 12/12/10</td>
<td>1 day</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>16 - Identification</td>
<td>Mon 12/13/10</td>
<td>Mon 12/13/10</td>
<td>0 days</td>
<td>100%</td>
<td>15</td>
</tr>
<tr>
<td>17 - Forensics</td>
<td>Mon 12/13/10</td>
<td>Thu 1/16/11</td>
<td>16 days</td>
<td>100%</td>
<td>16</td>
</tr>
<tr>
<td>18 - Systems Unavailable</td>
<td>Fri 1/7/11</td>
<td>Tue 1/11/11</td>
<td>3 days</td>
<td>100%</td>
<td>17</td>
</tr>
<tr>
<td>19 - Rebuild Systems</td>
<td>Fri 1/7/11</td>
<td>Tue 1/22/11</td>
<td>31 days</td>
<td>100%</td>
<td>17</td>
</tr>
<tr>
<td>20 - Rebuild Archive</td>
<td>Mon 1/13/11</td>
<td>Wed 3/16/11</td>
<td>32 days</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>21 - Final V6 Development and Checkout</td>
<td>Wed 1/15/11</td>
<td>Wed 9/28/11</td>
<td>185 days</td>
<td>13%</td>
<td></td>
</tr>
<tr>
<td>22 - L2 PGié Testing - Startup Status</td>
<td>Fri 2/25/11</td>
<td>Thu 4/28/11</td>
<td>45 days</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>23 - Match-Up and Cal Subset</td>
<td>Fri 2/25/11</td>
<td>Thu 3/24/11</td>
<td>20 days</td>
<td>100%</td>
<td>12F S-10 days</td>
</tr>
<tr>
<td>24 - Generate L2 Products (extended period)</td>
<td>Fri 2/25/11</td>
<td>Mon 4/11/11</td>
<td>32 days</td>
<td>100%</td>
<td>12F S-10 days</td>
</tr>
<tr>
<td>25 - Science Team Net meeting</td>
<td>Thu 4/14/11</td>
<td>Thu 4/14/11</td>
<td>1 day</td>
<td>100%</td>
<td>23, 24</td>
</tr>
<tr>
<td>26 - Science Team Meeting</td>
<td>Tue 4/19/11</td>
<td>Tue 4/22/11</td>
<td>3 days</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>27 - Post Science Team Meeting Activities</td>
<td>Fri 4/29/11</td>
<td>Wed 7/27/11</td>
<td>61 days</td>
<td>0%</td>
<td>26</td>
</tr>
<tr>
<td>28 - Final L2 Mods</td>
<td>Fri 4/29/11</td>
<td>Thu 6/30/11</td>
<td>44 days</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>29 - JPL Testing</td>
<td>Fri 6/10/11</td>
<td>Wed 7/27/11</td>
<td>32 days</td>
<td>0%</td>
<td>28F S-15 days</td>
</tr>
<tr>
<td>30 - Science Team V6 Checkout</td>
<td>Mon 6/27/11</td>
<td>Wed 7/27/11</td>
<td>21 days</td>
<td>0%</td>
<td>28F S-21 days</td>
</tr>
<tr>
<td>31 - Level 3 Algorithm Development</td>
<td>Wed 1/5/11</td>
<td>Mon 8/14/11</td>
<td>144 days</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>32 - L3 Concept Development</td>
<td>Wed 1/5/11</td>
<td>Wed 5/11/11</td>
<td>89 days</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>33 - L3 Coding</td>
<td>Thu 6/12/11</td>
<td>Mon 7/25/11</td>
<td>56 days</td>
<td>0%</td>
<td>32</td>
</tr>
<tr>
<td>34 - L3 Testing</td>
<td>Tue 7/12/11</td>
<td>Mon 8/11/11</td>
<td>15 days</td>
<td>0%</td>
<td>33, 34</td>
</tr>
<tr>
<td>35 - Package and Deliver</td>
<td>Tue 8/2/11</td>
<td>Tue 8/23/11</td>
<td>16 days</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>36 - Delivery Readiness Review</td>
<td>Tue 8/2/11</td>
<td>Tue 8/2/11</td>
<td>1 day</td>
<td>0%</td>
<td>36, 37, 38, 39, 40</td>
</tr>
<tr>
<td>37 - V6 Delivery Build</td>
<td>Wed 8/3/11</td>
<td>Wed 8/3/11</td>
<td>1 day</td>
<td>0%</td>
<td>39</td>
</tr>
<tr>
<td>38 - Final Checkout</td>
<td>Thu 8/4/11</td>
<td>Thu 8/5/11</td>
<td>4 days</td>
<td>0%</td>
<td>37</td>
</tr>
<tr>
<td>39 - V6 Packaging</td>
<td>Wed 8/10/11</td>
<td>Tue 8/16/11</td>
<td>5 days</td>
<td>0%</td>
<td>38</td>
</tr>
<tr>
<td>40 - V6 Code Delivery</td>
<td>Wed 8/11/11</td>
<td>Thu 8/18/11</td>
<td>2 days</td>
<td>0%</td>
<td>39</td>
</tr>
<tr>
<td>41 - JPL Unpack and Checkout</td>
<td>Fri 8/19/11</td>
<td>Tue 8/23/11</td>
<td>3 days</td>
<td>0%</td>
<td>40</td>
</tr>
<tr>
<td>42 - Hand-off to GES DISC</td>
<td>Tue 8/23/11</td>
<td>Tue 8/23/11</td>
<td>0 days</td>
<td>0%</td>
<td>41</td>
</tr>
<tr>
<td>43 - Documentation</td>
<td>Tue 7/5/11</td>
<td>Fri 9/16/11</td>
<td>53 days</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>44 - Documentation - all testing</td>
<td>Tue 7/5/11</td>
<td>Fri 9/16/11</td>
<td>53 days</td>
<td>0%</td>
<td>43</td>
</tr>
<tr>
<td>45 - Deliver Documentation</td>
<td>Fri 9/16/11</td>
<td>Fri 9/16/11</td>
<td>0 days</td>
<td>0%</td>
<td>44</td>
</tr>
<tr>
<td>46 - GES DISC I&amp;T</td>
<td>Wed 8/24/11</td>
<td>Wed 9/28/11</td>
<td>25 days</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>47 - V6 Integration and Test (GES DISC)</td>
<td>Wed 8/24/11</td>
<td>Wed 9/14/11</td>
<td>15 days</td>
<td>0%</td>
<td>47</td>
</tr>
<tr>
<td>48 - V6 End-to-End Testing</td>
<td>Wed 9/15/11</td>
<td>Tue 9/27/11</td>
<td>9 days</td>
<td>0%</td>
<td>47</td>
</tr>
<tr>
<td>49 - V6 Data Processing Begins</td>
<td>Wed 9/28/11</td>
<td>Wed 9/28/11</td>
<td>1 day</td>
<td>0%</td>
<td>48</td>
</tr>
<tr>
<td>50 - V6 Product Support Begins</td>
<td>Wed 9/28/11</td>
<td>Wed 9/28/11</td>
<td>0 days</td>
<td>0%</td>
<td>49</td>
</tr>
</tbody>
</table>
Upcoming V6 Milestones

- **V6 Concluding Timeline**
  - **JUN 30** - Final “final” L2 V6 coding/mods incorporated
  - **JUL 25** - Final L3 coding mods
  - **Mid-JUN** - Testing and checkout begins (incl. SciTeam)
  - **AUG 1** - Testing/checkout ends
  - **AUG 9** - Final Build and checkout
  - **AUG 23** - Hand-off to GES DISC, code and documents
  - **SEP 28** - Public Release V6 Data Products

- **This schedule allows for some minimal adjustments to be made during the testing period. Schedule risk is involved!**